

EMNLP 2017

**The First Workshop on
Speech-Centric Natural Language Processing**

Proceedings of the Workshop

September 7, 2017
Copenhagen, Denmark

©2017 The Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
209 N. Eighth Street
Stroudsburg, PA 18360
USA
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

ISBN 978-1-945626-92-0

Introduction

The first Workshop on Speech-Centric Natural Language Processing (SCNLP 2017) took place on Thursday, September 7, 2017 in Copenhagen, Denmark, immediately preceding the Conference on Empirical Methods in Natural Language Processing (EMNLP).

The purpose of this workshop was to unite the automatic speech recognition (ASR) and natural language processing (NLP) communities to discuss new frameworks for exploiting the rich information present in the speech signal to improve the capabilities of natural language processing applications. Our community objective is to revisit the conventional NLP problems with a focus on incorporating the richness of spoken language, as well as to encourage research contributions that promote cross-fertilization between statistical methods for ASR and NLP.

Our inaugural workshop was held at EMNLP to encourage participation amongst the NLP community to consider and discuss the challenges of combining speech recognition with conventional NLP research, as well as to appreciate the recent successes in this exciting field. The authors in these proceedings have combined ASR and NLP in works that address part-of-speech tagging, constituency parsing and dependency parsing on speech, information extraction and spoken term detection, dialog state tracking and speech translation, as well as two research assessments that evaluate the fluency and adequacy of English speakers and the role of speech silence in conversational dialogs.

The invited talk was given by Gabriel Skantze, entitled “Modelling turn-taking in spoken interaction.”

Our workshop also contained an open round-table discussion about the current state of speech-centric NLP and some of the research and pragmatic issues that raise a barrier of entry for the larger research community.

We would like to thank the members of the Program Committee for their reviews, as well as our panelists who led our round-table discussion. We also would like to thank the authors for their contributions.

Nicholas Ruiz and Srinivas Bangalore
Co-organizers

Organizers:

Nicholas Ruiz, Interactions (USA)
Srinivas Bangalore, Interactions (USA)

Program Committee:

Francisco Casacuberta, Universitat Politècnica de València (Spain)
Eric Fosler-Lussier, The Ohio State University (USA)
Dilek Hakkani-Tür, Google (USA)
Xiaodong He, Microsoft Research (USA)
Peter Heeman, Oregon Health & Science University (USA)
Julia Hirschberg, Columbia University (USA)
Preethi Jyothi, IIT Bombay (India)
Gakuto Kurata, IBM Research, Tokyo (Japan)
Lin-shan Lee, National Taiwan University
Yang Liu, University of Texas at Dallas (USA)
Karen Livescu, Toyota Technological Institute at Chicago (USA)
Raymond Mooney, University of Texas at Austin (USA)
Satoshi Nakamura, Nara Institute of Science and Technology (Japan)
Mari Ostendorf, University of Washington (USA)
Giuseppe Riccardi, University of Trento (Italy)
Andrew Rosenberg, IBM T.J. Watson Research Center (USA)
Isabel Trancoso, Laboratório de Sistemas de Língua Falada (Portugal)
Jason Williams, Microsoft Research (USA)

Invited Speaker:

Gabriel Skantze, KTH Royal Institute of Technology (Sweden)

Table of Contents

<i>Functions of Silences towards Information Flow in Spoken Conversation</i> Shammur Absar Chowdhury, Evgeny Stepanov, Morena Danieli and Giuseppe Riccardi	1
<i>Encoding Word Confusion Networks with Recurrent Neural Networks for Dialog State Tracking</i> Glorianna Jagfeld and Ngoc Thang Vu	10
<i>Analyzing Human and Machine Performance In Resolving Ambiguous Spoken Sentences</i> Hussein Ghaly and Michael Mandel	18
<i>Parsing transcripts of speech</i> Andrew Caines, Michael McCarthy and Paula Buttery	27
<i>Enriching ASR Lattices with POS Tags for Dependency Parsing</i> Moritz Stiefel and Ngoc Thang Vu	37
<i>End-to-End Information Extraction without Token-Level Supervision</i> Rasmus Berg Palm, Dirk Hovy, Florian Laws and Ole Winther	48
<i>Spoken Term Discovery for Language Documentation using Translations</i> Antonios Anastasopoulos, Sameer Bansal, David Chiang, Sharon Goldwater and Adam Lopez .	53
<i>Amharic-English Speech Translation in Tourism Domain</i> Michael Melese, Laurent Besacier and Million Meshesha	59
<i>Speech- and Text-driven Features for Automated Scoring of English Speaking Tasks</i> Anastassia Loukina, Nitin Madnani and Aoife Cahill	67
<i>Improving coreference resolution with automatically predicted prosodic information</i> Ina Roesiger, Sabrina Stehwien, Arndt Riester and Ngoc Thang Vu	78

Conference Program

September 7, 2017

8:50–9:00 *Opening Remarks*
Nicholas Ruiz and Srinivas Bangalore

9:00–10:00 Invited Talk

9:00–10:00 *Modelling turn-taking in spoken interaction*
Gabriel Skantze

10:00–10:30 Session I

Functions of Silences towards Information Flow in Spoken Conversation
Shammur Absar Chowdhury, Evgeny Stepanov, Morena Danieli and Giuseppe Riccardi

10:30–11:00 Coffee Break

11:00–12:30 Session II

Encoding Word Confusion Networks with Recurrent Neural Networks for Dialog State Tracking
Glorianna Jagfeld and Ngoc Thang Vu

Analyzing Human and Machine Performance In Resolving Ambiguous Spoken Sentences
Hussein Ghaly and Michael Mandel

Parsing transcripts of speech
Andrew Caines, Michael McCarthy and Paula Buttery

Enriching ASR Lattices with POS Tags for Dependency Parsing
Moritz Stiefel and Ngoc Thang Vu

12:30–14:00 Lunch

September 7, 2017 (continued)

14:00–15:30 Session III

End-to-End Information Extraction without Token-Level Supervision

Rasmus Berg Palm, Dirk Hovy, Florian Laws and Ole Winther

Spoken Term Discovery for Language Documentation using Translations

Antonios Anastasopoulos, Sameer Bansal, David Chiang, Sharon Goldwater and Adam Lopez

Amharic-English Speech Translation in Tourism Domain

Michael Melese, Laurent Besacier and Million Meshesha

Speech- and Text-driven Features for Automated Scoring of English Speaking Tasks

Anastassia Loukina, Nitin Madnani and Aoife Cahill

15:30–16:00 Coffee Break / Poster Discussion

16:00–16:25 Session IV

Improving coreference resolution with automatically predicted prosodic information

Ina Roesiger, Sabrina Stehwien, Arndt Riester and Ngoc Thang Vu

16:25–17:50 Round-table: Issues in Speech-centric NLP

17:50–18:00 Closing